The word and syntax

Words as grammatical units

Nikolas Gisborne, University of Edinburgh

n.gisborne@ed.ac.uk

[Author's final version; Chapter 11 of *The Oxford Handbook of the Word*, edited by John R. Taylor, Oxford: OUP. 2015.]

Abstract

The identification of what words are is central to grammatical theorizing: unless there is a theory of the word, it is impossible to establish what it is that syntax combines and what semantics composes the meanings of. But there is a substantial degree of theoretical debate about where the boundaries of words are and how the relationship between syntax and morphology should be understood. This chapter addresses that literature by exploring a case of inflection in English, the passive participle, and two "boundary" topics: cliticization and Noun Incorporation. It is argued that the Strong Lexicalist Hypothesis—the claim that neither inflectional and derivational morphology is syntactic—is adequate in these cases. The paper concludes that the Word is a theoretical construct, and that we should not expect naïve definitions to deliver useful analytic results. It is further argued that discourse reference and other semantic phenomena are not relevant to establishing wordhood.

Keywords

Lexicalism, inflection, derivation, clitics, noun incorporation, discourse reference, affixes.

1. Introduction

Most modern theories of syntax, including Generalized Phrase Structure Grammar (GPSG: Gazdar, Klein, Pullum, and Sag 1985), Head-driven Phrase Structure Grammar (Pollard and Sag 1994), Lexical-Functional Grammar (LFG: Bresnan 2001), and Minimalism (Chomsky 1995) as well as less familiar theories such as Word Grammar (WG: Hudson 2007), are lexicalist. Lexicalist theories treat words as the atoms of syntax, so they adopt what has been called the "strong lexicalist hypothesis" (Halle 1973, Lapointe 1980, Di Sciullo and Williams 1987). This hypothesis states that the job of

syntax is to combine words, or to state well-formedness constraints on combinations of words. The creation of inflected words or new words through derivation or compounding belongs elsewhere, in morphology. There is a weaker theory (the "weak lexicalist hypothesis": Aronoff 1976, Anderson 1982) which holds that compounding and derivation belong in morphology, but that inflection is syntactic. Among modern theories of grammar, there is one major school which rejects lexicalism: Distributed Morphology (Halle and Marantz 1993, Embick and Noyer 2007, Embick and Marantz 2008). The lexicalist turn in syntax follows Chomsky (1970). In this chapter, I address what it means to assert that the word is a "syntactic atom" (Ackema and Neeleman 2002).

In any discussion of the word and its relationship to syntax, there are potential confounds in the terminology. For example, some authors talk of inflection and derivation being "in the lexicon" and the combination of words into phrases being "in the syntax". This follows from a temporal metaphor which assumes a procedural architecture of grammar where syntax is derivational, and where morphology "takes place" in the lexicon before a word enters the syntax and potentially participates in derivational processes. I find the terminology unhelpful, because it often leads to a conflation of morphology with the lexicon and because theories such as HPSG, LFG, and WG reject derivations in favour of constraint satisfaction.

The lexicon includes information about words that is not relevant to the issue of whether words or subparts of words are the atoms of syntax: their meanings, for example. In addressing the question of what it means for a theory to be lexicalist, we need to understand that the debate is just this: does syntax need to look at subparts of words or not? What is at stake is whether the better (simpler, more psychologically plausible) theory is one where the same system creates new words, inflected words, phrases, clauses and sentences or whether it is one where there are discrete systems for creating words and for combining them.

In terms of linguistic theorizing, a lot hangs on this debate. There are two main reasons. One is in the nature of the complexity of the data: a substantial degree of the complexity of natural language belongs at the interfaces between subsystems, and the main task of the linguist is to find the regularities and generalizations. Phenomena such as "clitics" and Noun Incorporation, discussed in section 3, present significant

descriptive problems posing major challenges for linguistic theories.[1] The other main reason is that natural language is a complex system, and therefore so too are the theories that set out to describe it. There are consequences to adopting a particular position on the modularity or contiguity of syntax with respect to morphology. For example, Baker (1988, 1996) treats Noun Incorporation as syntactic and his work presents the original arguments for the theory of head movement. If NI turns out not to be syntactic, and therefore not to involve head movement, how much motivation is left for the theory of head movement?

In this chapter I focus on three topics, presenting and defending the lexicalist position in each case. The first case study section 2) looks at inflection—the passive participle. Section 3 looks at two phenomena at the boundary of syntax and morphology: French pronominal affixes (typically treated as clitics) and the phenomenon of clitic "climbing", and Noun Incorporation, especially the Mohawk data which has been generally taken to be particularly problematic for the lexicalist position. In section 4, I conclude and explore some wider perspectives on the atoms of syntax.[2]

2. Evidence from the passive participle

There are two kinds of evidence from the passive participle. One is to do with its relationship to word formation; the other with its distribution. Both of these concern the relationship of passive participles to other kinds of participle. It is possible for a participle of any sub-type to be converted into an adjective. Although there are historically lexicalized forms, such as CHARMING and INTERESTING, there are also nonce conversions.

2.1 Passive participles and word formation

Chomsky (1957) presented a syntactic treatment of passivization which involved deriving the passive participle in the syntax—a syntactic approach to inflection.

[1] I have "clitics" in scare quotes because this is a descriptive term capturing a range of phenomena which are not amenable to a single formal treatment. See Zwicky (1977, 1995).

[2] This chapter takes a syntactic view of lexicalism. There is a substantial morphological literature which looks at the same general question, but from the point of view of morphology. Lieber and Scalise (2007) review the history of the lexical integrity hypothesis and explore various morphological phenomena which pose challenges to it, including, for example, phrasal compounds in various Germanic languages.

Subsequent developments have led to the passive participle being treated as morphological. The weak lexicalist position is that derivation is morphological, but inflection is syntactic. But what if inflected forms are the input to a word formation process? If that were the case, then both inflection and word formation must be morphological, and neither could be syntactic. Bresnan (1978, 1982) has argued just this; her arguments are summarised in Bresnan (2001: 30-39). The claim is that passive participles have two distributions, as non-finite verbs and as adjectives; the adjectival passive participles are derived from the verbal; the verbal passive participles are an inflected form of the verb; and therefore inflection comes before derivation, and both must take place in the same domain of the grammar: morphology.

First, she shows that participles in general convert to adjectives. The examples in (1) below are Bresnan's (2001: 31) example (10).

(1) a. *present participles*: a smiling child, a breathing woman, the boring story
b. *perfect participles*: a fallen leaf, an escaped convict, wilted lettuce
c. *passives*: a considered statement, the spared prisoners, an opened can

These converted participles have the properties of adjectives: it is possible to have negative (rather than reversive) *un*-prefixation; they are gradable; and they can head concessive relatives beginning with HOWEVER.

(2) an unconsidered statement; a very considered statement; however considered her statement might have been

The full range of passive participial forms counts for conversion to adjectives, as in (3). The examples are Bresnan's (2001: 31) example (12).

(3)

Verb	*Participle*	*Adjectival Participle*
sing	sung	an unsung hero
fight	fought	hard-fought battles
write	written	a well-written novel
give	given	a recently given talk
consider	considered	an unconsidered action

inhabit	inhabited	an uninhabited island
break	broken	my broken heart
split	split	split wood

The examples in (3) show that the same form occurs as verbal passive participle and as adjective passive participle. Therefore, adjectival passives must have verbal passives as their input, otherwise there is no way to account for the morphological parallels.

Another argument comes from prepositional passives: it is possible to have adjectival versions of prepositional passives, as in (4), and where it is not possible to have a verbal prepositional passive, the adjectival passive is likewise ruled out (5 & 6). Again the examples come from Bresnan (2001: 31-32).

(4) a. After the tornado, the fields had a *marched through* look.
b. Each *unpaid for* item will be returned.
c. You can ignore any recently *gone over* accounts.
d. His was not a *well-looked on* profession.
e. They shared an *unspoken, unheard of* passion for chocolates.
f. Filled with candy wrappers and crumpled bills, her bag always had a *rummaged around in* appearance.

The next two pairs of examples show that if a verb cannot have a prepositional passive then neither can there be an adjective derived from it which permits a prepositional passive.

(5) a. *The twin is looked like by his brother.
b. *a looked like twin (cf. like-minded)

(6) a. *No reason was left for.
b. *the left-for reason (cf. each unpaid-for item)

Given the above data above, the simplest story is that the verbal passive participle is an input to adjective formation. It makes far less sense to say that there is a separate process of passive adjective formation which runs parallel to a mechanism that derives

the passive construction. And for these reasons, then, Bresnan (1978, 1982) argued that passive was a morphological and not a syntactic rule: passive participles, and the adjectives derived from them, are formed in the lexicon.

Bresnan has shown that both kinds of passive have to be lexical while treating the verbal passive as inflection and the adjectival passive as word formation—which in turn provides evidence for the strong lexicalist hypothesis, or Zwicky's (1992: 354-5; 1996: 30) Principle of Morphology Free Syntax (see also Brown and Hippisley 2012).[3]

But there are still some loose ends. First, how do we account for prepositional passives such as *these woods were walked in by Anne Boleyn* in a lexicalist theory? And secondly, what about verbs which apparently have objects, but which do not figure in the passive construction, and which appear not to have passive participles, such as RESEMBLE, FIT, INVOLVE, WEIGH? I return to these in section 2.3. In the next section, I present a distributional argument in favour of the strong lexicalist position, that is, that the behaviour of verbal passive participles argues in favour of the lexicalist position.

2.2 The distribution of passive participles and other inflected elements

Although I am not aware of any published work that argues for a lexical account of passive participles from their distribution, although Hudson (1990: 339) has pointed out that passive participles and the *–ing* participle have the same distribution, it seems to me that one of the major arguments in favour of the lexicalist position is the distribution of lexical items, including passive participles. Passive participles can be found in a number of different environments. But before we discuss their distribution we need first to agree on terms. The distributional argument requires one additional premise: that forms realize cells in a paradigm.[4] I take the view that passive participles are exponents of a feature bundle which includes information such as the verb's name, and the fact that it is a verb, the fact that the verb has passive voice which is realized as VERB-*ed*. Because of the general property of syncretism, we can (and need to) distinguish

[3] Bresnan dealt a significant blow to Wasow's (1977) claim that adjectival passives had a different derivation from verbal passives, as well as demonstrating alongside Wasow the incorrectness of Friedin's (1975) claim that all passive participles were adjectives.

[4] It is possible for a theory to be lexicalist and morphemic or lexicalist and realizational (see Stump 2001: 1-3). The argument developed in this section requires us to adopt a theory which is both lexicalist and realizational.

between the passive participle and the other verbal morphosyntactic form which is realized as VERB-*ed*, the perfect participle. This common form is a syncretism, which is the property where one form occurs in more than one cell in a paradigm (Matthews (1991: 202).

We have to distinguish passive participles from perfect participles because they have different feature specifications. The form *broken* is three-ways ambiguous: it could be the passive participle (*my glasses were broken by some idiot*); the perfect participle (*you have broken my glasses*); or the adjective (*my glasses seemed broken*). The passive participle has the same distribution as the active participle ending in *ing*; the perfect participle is more restricted, and typically occurs as the predicative complement of HAVE. Neither the active nor the passive participle occur in that position. I show the basic distributions in (7); the other contexts that passive participles occur in are discussed below.

(7) a. The walls were repainted yesterday afternoon (by the new builders). [Passive]
b. The children were running in the 3-legged race yesterday afternoon. [Active]
c. The children have broken their ankles. [Perfect]

First of all, note that neither the passive not the active participle can occur as the predicative complement of HAVE. Note too that the perfect participle cannot occur as the predicative complement of BE.

(8) a. *The walls have repainted yesterday afternoon by the new builders.
b. *The children have running in the 3-legged race yesterday afternoon.
c. *The children are broken their ankles.

I have kept the voice features stable in (8a) and (8c): *repainted* has its patient argument as its subject, and its agent is expressed in the phrase *by the builders*; likewise, *broken* has a direct object: *their ankles*. From this, we have to conclude that the passive and the perfect participles are distributionally different as well as having different feature specifications. Their formal similarity is just a syncretism.

We can now turn to the range of distributions of the passive participle. I take the view that the passive participle itself is what has the distribution, not passive clauses

headed by that participle, but nothing much hinges on the distinction in the argument that follows. We can start with the GET passive in (9a,b) which looks very like the prototypical passive in (7a). The example in (9c) is not like a typical passive clause, because here the subject of *broken* is *your neck*, which is not the subject of *get*.

(9) a. His leg got broken in the car accident.
b. He finally got kissed by the girl he fancied.
c. Try not to get your neck broken when you're climbing that mountain.

As I said above, *broken* is three-ways ambiguous, and I used SEEM as a diagnostic of an adjective. Verbs like SEEM are not like BE and GET. The example in (10a) is adjectival not verbal, as (10b,c) show: SEEM is not a copular verb and cannot have a verbal predicative complement, although it can take an adjective (10d).

(10) a. We'd better not move him: his leg seems broken.
b. *His leg seemed broken by the lorry's wheel.
c. *The child seemed running fast.
d. The child seemed happy.

The use of the BY-phrase forces a long passive interpretation, and the example in (10b) is ungrammatical. In (10c), we see that SEEM cannot have an active participle as its predicative complement either, so the generalization is that SEEM is one of the verbs which does not select for participle predicative complements.

However, we do see passive verbs as the predicative complements of those other non-copular verbs which do permit verbal elements as their predicative complements. For example, passive participles can occur as the complement of causative and experiential HAVE and SEE.

(11) a. The drug-lord$_i$ had his enemy's$_j$ legs broken (by his$_i$ thugs).
b. She had her car stolen from her driveway last week.
c. We saw the city destroyed by enemy troops.

And there are constructions with COME and GO which also occur with passive clauses. These examples are from Pullum (2011).

(12) a. The problems with the building went unnoticed by the owners for weeks.
b. This software comes pre-installed by the manufacturers.

There is a construction with NEED which shows dialectal variation—in standard English, (13a) is the usual form; in Irish and Scottish English, it is (13b). The point of the example in (13a) is that in one particular construction, the property of having passive voice is exceptionally realized by the *–ing* participle rather than the *–ed* one.

(13) a. The windows need washing (='To be washed')
b. The windows need washed.
c. He likes counselling.

The example in (13c) is ambiguous: it can either be interpreted as he likes being counselled, or as he likes counselling others. I take it that this speaks to the potential ambiguity of the *–ing* form.

There are two further distributions which are relevant to the issue of where participles are created: they occur as the adjuncts of nouns, in participial relatives (14) and as clause initial adjuncts (15). In (14), the bracketed elements are Noun Phrases with participial relatives modifying the noun.

(14) a. [The student attacked by the police] successfully sued.
b. [The speaker acclaimed by the senator] was a dreadful drunk.

(15) a. Attacked by wolves, the cat raced up the tree.
b. Running home, the boy tripped over.

The clausal adjuncts in (15a) and (15b) are both participial constructions; the example in (15a) involves a verbal passive participle. It's not adjectival.

The first point to be established is that each of the environments where a passive participle can occur is also an environment where the active participle in *–ing* can occur. Therefore, not only is there no separate transformational rule for each of the different possible passive constructions, but also the distribution of passive participles does not even need to be stated as a constraint on **passives**: it is a constraint on **participles**. I come back to what that means for an account of the lexicalist hypothesis below, but first let us go through the examples.

We have seen that both active and passive participles can occur as the complement of BE. It is sometimes argued, in fact it is a very common textbook position, that there must be different auxiliary verbs BE, one which selects the passive participle and one which selects the active one. This claim is unwarranted, but it is argued for on the basis of auxiliary order in (16).

(16) He may have been being beaten.

Here, *have* selects the perfect participle *been*, which in turn selects the active participle *being*, which selects the passive participle *beaten.* However, that sequence of selection patterns is no reason to assume that there are two auxiliary verbs BE. In the case of auxiliary BE, all we have to say is that it selects a participle. Everything else follows. When the active participle is itself an instance of BE, then—because it's BE—it selects a participle. BE cannot have a passive participle because it does not have an object. And there are no constructional semantics: the semantics of VERB-*ing* follow compositionally, as do the semantics of VERB-*ed.*

The next case to be discussed is the complementation of GET. The examples in (17) show GET with VERB-*ing* structures, both with and without an object.

(17) a. He got cooking as soon as he arrived at home.
b. Try to get your children cooking early in life, so that they can look after themselves.

These patterns match those in (9). I have already shown that SEEM does not allow participle complements in (10) so I move on to causal and experiential HAVE and SEE.

(18) a. She had his heart racing whenever she came into the room.
b. He had his heart racing after 5 minutes on the treadmill.
c. We saw the dog running across the road.

In (18), the present participles could each be replaced by a passive participle and the distribution is just the same as in (11).

In the case of COME and GO, the interpretation is depictive, as it is with passive participles, and the structures admit active participles just as well as passive ones.

(19) a. The children went running past all afternoon.
b. I inherited an old Persian carpet; it came crawling with bugs.

Just for now, I am putting the examples with NEED to one side—they feature at a later stage in the argument. The next distributional pattern is the participial relative, and here too we can have active participles just as well as passive ones.

(20) a. [The student attacking the police] was a trot.
b. [The speaker droning on] bored even his fans.

And finally, we see active participles as clause-initial adjuncts.

(21) a. Running home, I tripped and fell.
b. Advancing on the enemy, he died in the first skirmish.

The first task is to establish the right generalization, which is that active and passive participles share their distribution. Therefore, we do not want to discuss the distribution of passive participles as a separate phenomenon. More to the point, although this has long-since been decided, there can be no passive transformation and passive clauses are not derived from active clauses.

However, the shared distribution does not necessarily get us to a lexicalist story. It could be argued that the distribution is shared by *–ing* and *–ed*$_1$ (where the subscript "1" distinguishes this *–ed* from the other one found in the perfect verb form). On such a theory, the ending is the head of the word, and the grammar can look "inside" the word

to see what its distribution is: parts of words are the terminal nodes of syntactic trees or dependencies.

There are two arguments against this position: the behaviour in (13) with NEED, and suppletion. In (13a), passive voice is expressed by a participle ending in *–ing*. How are we to understand this? According to the model I have been arguing for in this section, I would say that *washing* is an unanalysed whole, and that it is the contextually determined form which realizes the passive of WASH. If, on the other hand, we were to argue that there was a form, *-ing*, which had its own lexical entry, stating that it was a head, that its morphological root was its syntactic dependent, and that it was active voice, we would need a new *–ing* to capture the facts in (13a). This would cause us to set up two distinct lexical entries, active voice $–ing_1$ and passive voice $–ing_2$. It is certainly possible to do that, and I own that there are people who might want to make such a move, but I think that the alternative position is better.

The alternative position is the realizational argument I offered at the beginning of this section. According to this position, the *–ing* form in (13a) is just a lexically conditioned realization of the passive participle. This analysis has some properties which are simpler than the alternative one, with multiple entries in the lexicon for *–ing* and *–ed*, as well as an entry for *–en* which we would also need to capture examples such as *broken*. First of all, it reduces the number of forms-with-meanings in the lexicon: realizations do not have meanings themselves; the paradigm cells they realize do. Secondly, it means that the dialectal fact, where in one variety *–ing* can realize either the active or the passive participle, is just a local irregularity in one of the cells of a paradigm. No new word needs to be set up to capture that fact and we know from other dialects of English that there can be quite a lot of variation in the realizational facts of English verb morphology. For example, there are northern English varieties that have *were* as the past tense of BE through the system, singular and plural. Likewise, there are varieties from the south west of England that have *be*.

Why might it be better to locate this variation in a theory of realizations rather than a theory that says there are different word-entries across the different dialects? One reason is that you should expect word entries to be more stable than realizations, because word entries are linked to more information. They are related to the word class of the form, and also to its meaning. A realization, on the other hand, is related to the cell in the paradigm that it realizes. The cell in the paradigm does the work of relating

the word, its class, and its meaning together. So in a realizational theory, parts of words are not themselves meaningful. Matthews (1991: 180) presents a clear case of the difficulty of a morphemic analysis, and working out what means what, in his treatment of Ancient Greek *elelýkete*; he also points out that the boundaries of morphological forms can be hard to establish (1991: 203). But there is a simpler argument.

When he was aged between four and five, my younger son came out with the utterance in (22).

(22) *I wish I didn't hadded a big brother. I wish I hadded a dog.*

What is important about this example is that the semantics is really difficult: my son had to know that under a verb like WISH he's expressing irrealis semantics. He also had to know that in subordinate clauses, irrealis semantics is expressed by a past tense verb. In the example, these two things are perfectly expressed. What he didn't know was how to express past tense. The irrealis semantics is right; the use of past tense is right; but the realization of past tense is a complete mess: he's got triple tense marking in the first clause, and double in the second. What does this mean? It must be an argument against a direct relationship between morphological forms and their meanings. If the forms were directly related to the meanings, then my son could not have got the semantics right while getting the morphology wrong.

And what of suppletion? This too is an argument for a lexicalist and realizational approach. There are suppletive forms for the passive participle which would make it impossible to build a morphemic analysis around *–ed*, and there are also forms replacing the *–ing* form. Both examples are non-standard, with (23b) perhaps being dialectal; (23c), however, is naturally occurring from an SMS/text message.

(23) a. The train has already gone.
b. It was broke when I got there.
c. I am sat on the bench in front of the station.

In (23a), *gone* is the lexically specific suppletive form of the perfect participle of GO. It's monomorphemic, so it's clearly not built around any of the common morphological

parts.[5] In (23b), *broke* realizes the passive voice of the verb BREAK. These kinds of suppletion are very common in non-standard Englishes. For example, where I live in Scotland, I often hear *the train has went* for *the train has gone* where the suppletive form is syncretic with the past tense form. The final example in (23c) is not passive but progressive and means the same as *I am sitting on the bench in front of the station*. Although it is "non-standard", it is common enough in the speech of speakers of Standard British English. Suppletion is straightforward to negotiate in a realizational theory: the realization of a given cell in a paradigm is not the default form. Rules which are more specific than defaults override defaults, and so the suppletive form is the realization of that cell across the grammars of a number of speakers.

So much for participles. What about the other forms of verbs? In brief, the same arguments apply, as I suggested when I discussed the dialectal forms of BE above. I take the view that the category "verb" does not supply any distributional information. It is the inflected forms of verbs that have distributions, which applies to tensed verbs every bit as much as to the participles. It is heads that select their dependents, and so finite verbs may be dependents only of those predicates that select them. Otherwise, they are the matrix verb in main predications; the lack of categorial distributional information on verbs is what allows gerunds to have the distribution of a noun and the complementation of a verb in a mixed category analysis (Hudson 2003; Malouf 1998), although Hudson claims that it is "non-finite verbs" that lack distributional information.

2.3 Prepositional passives and lexicalism

I now address the problems posed for the lexicalist account of passivization by cases such as prepositional passives, as in (24).

(24) a. This bridge has been flown under.
b. This paper has been written on both sides of.
c. The roof has been walked on.

There is an argument that asks how examples like (24) can be reconciled with a non-transformational account of passive: if the pair of words "FLY" and "UNDER" are not

[5] Historically, it is {go}+{en} but I take it that it is monomorphemic in the contemporary native speaker's mental lexicon.

stored together in the lexicon, how is a string like (24a) possible (McCawley 1998: 85-94, Bruening fc: 9)? Bresnan's lexical approach (1982) assumes reanalysis, with the passive participle and the preposition making a new complex predicate (see also Levin and Rappaport 1986). This reanalysis is assumed to be morphological, in that it creates a new item in the lexicon in that it is a kind of incorporation.[6]

However, Postal (1985), Baltin and Postal (1996), and Lødrup (1991) all find problems with Bresnan's analysis, summarized in Alsina (2009) which argues instead for structure sharing. The main problem for the lexical analysis that Postal (1996) and Baltin and Postal (1996) find is that the Verb+Preposition sequence does not behave like a word, but like a syntactic sequence. Alsina's proposal addresses the criticism by taking the relationship between the passive participle and the preposition to be syntagmatic and related to the phenomena of subject-to-subject raising and extraction.

As Huddleston and Pullum (2002: 1444) point out, there are discourse, semantic, and pragmatic constraints on passives. One of the discourse constraints is that the passive subject has to be discourse old, at least relative to the content of the passive BY-phrase. Hypothetically, any verb could have a passive participle, irrespective of whether it is transitive or not, just as any verb can have an *–ing* participle whether it is stative or not. For a passive participle to be used grammatically, either the subject of that participle must be in the semantic relation normally associated with its active voice object, or it must denote some entity which is affected by the action denoted by the verb, and which is present in the grammatical structure. Thus the grammaticality of the examples in (24)—we can treat flying under a bridge as affecting it in some way; likewise writing on both sides of a piece of paper. Any treatment, lexicalist or not, has to be sensitive to the discourse and semantic facts.

However, some verbs patently do not have passive participles; for a subset, this is because they do not have semantic arguments which map to their subjects, and so they do not license the basic requirement of a passive participle, which is that the semantic argument mapping to Subject in the active voice is delinked from the Subject in the passive voice. Such examples would include SEEM and WEIGH, where SEEM is a raising verb which does not collocate with prepositions at all, and WEIGH is a verb which takes a nominal predicative complement, with the meaning 'is, by weight'.

[6] In an alternative lexicalist account, Hudson (1990) analyses different subtypes of prepositional passive using a novel dependency-type, "passive-link". If dependencies are constructions (Gisborne 2008), then this is a kind of constructional analysis; it is certainly *ad constructionem*.

It is now possible to account for the adjectival passive of examples such as (24): the verb has a passive participle; therefore the passive participle is available for conversion into an adjective. The converted adjective retains the collocational possibilities of the verb it derives from because both verbs and adjectives can collocated with prepositions. When a verb has a PP complement, the derived adjective simply inherits that PP complement.

2.4 Evaluating the facts about passive

There are three different sets of data that I have drawn on in this section. The first two, the derivation of adjectival passives and the distribution of passive participles offer clear arguments in favour of a lexical treatment of passive participles. In particular, when there are clearly passive structures which are not derivable from active clauses, it is axiomatic that passive participles are formed in the morphology, and that passive constructions work by simple pattern matching. It follows that inflection is lexical.

There are adequate lexicalist accounts of the more complex data surrounding prepositional passives, so there is no need to invoke a derivational approach to account for them. One of the main points that follows from the discussion of prepositional passives is that it is specific proposals that need to be evaluated, not the general research strategy.

The main remaining issue is whether there need be any direct relation between active and passive clauses. I have argued that there doesn't, the main reason being that there are so many subtypes of passive clause that it makes no sense to privilege a particular case, although the mapping between active and passive voice motivates Collins (2005).

3. The boundaries of grammar

So far, I have argued that it is both possible and desirable to take a lexicalist stance on verb inflection and shown that this is necessary from the point of view of modern lexicalist theories of syntax. In developing the argument, particularly the distributional argument for the lexicalist position on the English passive, I argued for a realizational theory of morphology. Now I turn to two further sets of data which have been claimed to be problematic for the lexicalist position. The first problem is the behaviour of clitics, in particular clitic climbing. The second is Noun Incorporation. In both cases, it has been

shown that it is possible to handle these data without recourse to syntactic solutions, including "mixed" analyses, which treat certain kinds of phenomena as partly in one system and partly in another.

Clitics are a general research problem. As Zwicky (1977, 1995) pointed out, backed up in more depth by Miller (1992), Anderson (2005), and Spencer and Luis (2012), there are several different phenomena that have been identified as "clitics", and they are not amenable to a single analysis. Crudely speaking, clitics are elements that have the form of an affix, but the syntax of a word. They are therefore a problem for lexicalism, because they appear to be on both sides of the syntax-morphology boundary. One particular problem is the case of the pronominal affixes which I address in §3.1. These look very much like affixes, but they also display the properties of "clitic climbing" where a pronominal form is realized on a higher verb than the one whose argument structure it belongs to. To explore the issues, I discuss a particular case study: Miller and Sag (1997).

Noun Incorporation (NI) is a kind of compounding, which has been subject to both lexicalist and syntactic analyses. On the face of it, compounding should be a simple matter of morphology, but there are complications in some of the data. In particular, Mohawk has been subject to syntactic analysis by Baker (1988, 1996) because it has the property of incorporating an argument which is definite and discourse referential. It is widely assumed that only words can have these semantic properties, and consequently NI has become a general testbed for lexicalist and syntacticist accounts. Mithun (1984, 1986), Rosen (1989), DiSciullo and Williams (1987) all take a lexicalist approach; Sadock (1991) and Baker (1988, 1996) take a syntactic one.[7]

3.1 French pronominal affixes

French pronominal affixes include examples such as those in (25) from Miller and Sag (1997)—their example (1). In the literature, these are usually called clitics; in fact, they are one of the subtypes of Zwicky's (1977) 'special' clitics, discussed at length in Anderson (2005), but following Miller and Sag's analysis I refer to them as 'pronominal affixes'.

[7] One complication in the NI literature is whether the phenomenon in West Greenlandic discussed by Sadock is the same as the phenomenon discussed by Mithun, Rosen, and Baker. Noun Incorporation is arguably a sub-species of compounding; the West Greenlandic facts are more like derivational morphology.

(25) a. Marie le voit. 'Marie sees him.'
b. *Marie le voit Jean. 'Marie sees Jean.'[8]
c. Marie voit Jean. 'Marie sees Jean.'

These examples show that a pronominal formative, such as *le* in (25a), reduces the valency of the verb, which is why (25b) is ungrammatical. It is also possible to have combinations of pronominal formatives as in (26)—Miller and Sag's (2)—and for pronominal formatives to reduce more than one complementation property.

(26) a. Marie lui donne le livre. 'Marie gives her the book.
b. Marie le lui donne. 'Marie gives it to her.'
c. * Marie lui donne un livre à Anne. 'Marie gives a book to Anne.'
d. *Marie le lui donne le livre. 'Marie gives her the book.'

As Miller and Sag (1997) point out, these cases have been analysed in various ways. Some scholars have argued for analyses involving movement from an argument position (Kayne 1969, 1975, 1991; Perlmutter 1970); others have argued for base-generated analyses (Rivas 1977; Jaeggli 1982) which Sportiche (1996) has criticized on the basis of examples such as those in (27), which are Miller and Sag's (3). In these examples, the pronominal affixes are realized on an auxiliary verb rather than on the main verb whose valency they reduce. The phenomenon is known as clitic climbing.

(27)	a. Mari l'a vu.	'Marie has seen him.'
		[*l'* – argument of *vu*]
	b. Le livre lui a été donné.	'The book has been given to him.'
		[*lui* – argument of *donné*]
	d. Pierre lui reste fidèle.	'Pierre remains faithful to him.'
		[*lui* – argument of *fidèle*]
	e. Marie en connaît la fin.	'Marie knows the end of it.'
		[*en* – argument of *fin*]

[8] Miller and Sag (1997: 574, fn 2) point out that examples such as (25b) "are grammatical with a pause before the final element (*Jean*), which we take to be indicative of a right dislocated (unbounded dependency) structure"

f. Marie le fait lire à Paul. 'Marie is making Paul read it.'
[*le* – argument of *lire*]

The question posed by the examples in (27) is how it is possible to have a base generated account of the French pronominal affixes when they are dislocated from the verbal heads whose argument arrays they satisfy. The dislocated pronominals look like they violate the lexicalist hypothesis, because they behave like clitics which have moved. There is a further set of issues, spelled out by Bonami and Boyé : because there are particular constraints on the co-occurrence of French pronominal elements, it is widely held that they can be best analysed in a templatic structure assuming seven position classes (Bonami and Boyé 2007: 293). This makes the situation worse, because if the pronouns in (27) have moved, it would mean that there are both morphological constraints and syntactic movement.

Following Miller (1992), Miller and Sag (1997) demonstrate that it is a straightforward matter to analyse French pronominal clitics lexically while at the same time capturing the facts in (27). They analyse the pronominal elements as lexically attached inflections not as postlexical clitics. Their criteria for treating the pronouns as inflectional affixes draw on Zwicky and Pullum's (1983) arguments that English *n't* is an affix. It is also important to note that Miller and Sag (1997: 576) assume that the pronominal affixes really are pronouns. They state, "agreement marker vs. pronoun status and affix vs. word status are two independent parameters", so the bound pronouns are "affixal (or 'incorporated') pronouns. The evidence for this is the absence of systematic doubling".

This is a challenging position to adopt, because it keeps a number of theoretical consequences and decisions in play. First though, it requires us to keep a number of factors clearly separate. In Miller and Sag's (1997) theory, what is at stake is the relationship between morphology and syntax: as far as the syntax is concerned, these pronominal elements do not exist. The combinatorics of the sentence only involve the larger words which include the words which themselves include the pronouns. As far as the morphology is concerned, however, these pronominal elements do exist and they are part of word structure. What is not at stake, on the other hand, is the relationship between syntax and semantics, or morphology and semantics. These pronominal elements are fully referential, and therefore for Miller and Sag (1997) can refer

independently. This is significant, because it tells us where their theory locates complexity.

Crudely speaking, complexity can be found at the interfaces. A simple illustration involves the raising/control distinction, although this is about the syntax-semantics and syntax-phonology interfaces and doesn't strictly involve morphology. The difference between *I expected her to go* and *I persuaded her to go* is that with *expected*, *her* is not a semantic argument of the matrix verb whereas with *persuaded*, *her* is. If reducing complexity at the syntax-semantics interface is what matters, the solution is to posit that there is an additional, phonologically unrealized word (PRO) which is the subject of *to go* when the non-finite clause complements *persuaded*. If, on the other hand, what matters is minimizing complexity between morphology and phonology, then a theory will allow *her* under *persuaded* to serve syntactically in both clauses, reducing the complexity in terms of having elements which are phonologically unrealized, but increasing the complexity in terms of the relationship between syntax and semantics, by allowing an item to receive more than one semantic role. The theory presented by Miller and Sag (1997) puts the complexity at the syntax/semantics interface, in order to minimize complexity between syntax and morphology.

Miller and Sag (1997) exploit seven criteria, presented in (28), to establish that the pronominal elements are affixes. The first three are taken directly from Zwicky and Pullum (1983: 503-504), and to these I have added Zwicky and Pullum's comments.

(28) a. Degree of selection with respect to the host: clitics can exhibit a low degree of selection with respect to their hosts, while affixes exhibit a high degree of selection with respect to their stems.

b. Arbitrary gaps in the set of combinations, more characteristic of affixed words than of clitic groups.

c. Morphophonological idiosyncrasies which are more characteristic of affixed words than of clitic groups.

d. Rigid and idiosyncratic ordering which is "typical of affixation rather than of cliticization. For instance the ordering of dative and accusative pronominal affixes in standard French depends on the persons of the affixes involved." (Miller and Sag 1997: 579)

e. Pronominal affixes undergo lexical phonological rules: "affix+stem units undergo lexical phonological rules, such as obligatory liaison of nasal consonants". (Miller and Sag 1997: 579)
f. Object affixes cannot have wide scope over coordination. *_Pierre les voit et écoute_, which would mean something like 'Pierre sees and hears them', is ungrammatical and Miller (1992) argues that this is strong evidence for the lexically attached status of these elements. (Miller and Sag 1997: 579-580)
g. Syntactic explanations for clitic ordering have failed. Miller and Sag cite a literature which has shown that it is not possible to provide a principled syntactic account of pronominal affix ordering.

Note that all of these diagnostics are concerned with morphosyntactic diagnostics for the word boundary. This is even true of (f) which means that object affixes can only be construed as the object of the verb that they are attached to, and not the "understood" object of another verb in a coordination. Miller and Sag (1997) are therefore focused on establishing the limits of syntax with respect to word-boundaries. Their theory does not elaborate a wider notion of the word, perhaps with respect to a theory of signs, or with respect to a larger "constructicon". This point will be important in section 4.

Once they have made their empirical arguments about the nature of pronominal affixes, the crucial task for Miller and Sag is to account for the data in (27) above, especially the non-canonical placement of these affixes. Miller and Sag set themselves a particular challenge in that having demonstrated that French pronominal affixes are affixes, their realization on an auxiliary has to be accounted for.

Miller and Sag's solution exploits the logic of default inheritance and the possibility of multiple inheritance, as expressed in the architecture of HPSG. They argue (1997: 586) that there are two types of verbal realization: verbs can be realized as plain words, whose arguments are phrases, and cliticized words, which are verbs that have at least one argument that is realized affixally. They elaborate a theory that assumes that French auxiliaries and main verbs enter into a flat structure rather than a hierarchical one, with constituency tests confirming the analysis. This allows them to assume a lexical entry for the tense auxiliary verbs where the elements of the Argument Structure list of the participial complement are also arguments of the auxiliary. As a result, "the

argument structure of the participle determines that of the auxiliary verb's lexeme, and hence the valence of words formed from that auxiliary" (Miller and Sag 1997: 601). Therefore, they provide an argument-sharing analysis which "interacts with the analysis of pronominal affixation […] so as to predict a wide range of nonlocal pronominal affixation phenomena" (ibid). As a result, "what appears to be 'climbing' of clitic elements follows immediately from lexical principles and local syntactic combination, without the introduction of any further devices" (Miller and Sag 1997: 603).

In the next section, I show that the property of being independently referential has been criterial in some aspects of theorizing about incorporation. Miller and Sag, however, simply establish an affix type (either a-aff for an anaphoric affixal pronoun, or p-aff for a personal affixal pronoun) in their SYNSEM feature matrix. Miller and Sag's (1997) proposal gets the relevant grammaticality differences of the examples in (29).

(29) a. Jean$_i$ sait que Paul$_j$ s'aime. 'Jean$_i$ knows that Paul$_j$ loves himself$_{j,*i}$'
b. Jean$_i$ sait que Paul$_j$ l'aime. 'Jean knows that Paul$_j$ loves him$_{i,*j}$.'

As a result, the full referentiality of the pronominal affix is not an issue for them.

Miller and Sag's (1997) theory has a number of consequences for the theory of grammar. As we shall see in the next section, many have argued for a lexicalist theory of incorporation which similarly has broad consequences for the theory of grammar. Miller and Sag (1997: 633) raise doubt about the Head Movement Constraint, as Anderson (2000) does, and they also argue that the Mirror Principle—which holds that the ordering of morphemes mirrors that of the hierarchy of functional projection—is untenable. Ackema and Neeleman (2004) point out that in derivational syntax is it possible to create complex heads which have different properties from compounds in the same languages. See also Ackema and Neeleman (2010: 23).

3.2 Noun Incorporation

Noun Incorporation (NI) presents similar problems to French pronominal affixes. In NI, which is a type of compounding (Mithun 1984, Di Sciullo and Williams 1987, Rosen 1989, Giegerich 2009), an argument of the verb is realized within a morphologically complex verb as a nominal root; there are different types of NI, but in at least one subtype the incorporation of the nominal root reduces the verb's valency so that if the

incorporated nominal root corresponds to the verb's direct object it is not possible for there to be a discrete Noun Phrase direct object. The alternative analysis is that NI involves movement from the direct object position to the incorporation site/landing site within the complex verb (Baker 1988, 1996). NI is therefore a classic test-bed for questions of wordhood and theories of the word. Baker (1988, 1996) presents the main arguments for a syntactic treatment of NI; Anderson (2000, 2001, 2005: 257-287) usefully summarizes and presents a number of arguments for the lexicalist position.

The examples in (30) and (31) below are from Rosen (1989: 295), and are her examples (1) and (2).

(30) Onondaga (Woodbury 1975: 10)

a. *wa ʔhahninúʔ* *neʔ* *oyɛ́kwaʔ*.
TNS.3SG.3N.BUY.ASP ART 3.N.tobacco.NM
'He bought the tobacco'

b. *wa ʔha y ɛ ʔkwahní:nuʔ*.
TNS.3SG.3N.tobacco.buy.ASP
'He bought tobacco'

(31) Niuean (Seiter 1980: 69)

a. *Takafaga* *tu:mau* *ni:* *e* *ia* *e* *tau* *ika.*
hunt always EMPH ERG he ABS PL fish
'He's always hunting fish' (= He's always fishing)

b. *Takafaga* *ika* *tu:mauni:* *ni:* *a* *ia.*
hunt fish always EMPH ABS he
'He's always fish-hunting' (= He's always fishing)

There are different kinds of Noun Incorporating languages. In the Niuean example, the incorporated noun is non-specific and indefinite or generic. But in Mohawk, it is possible for the incorporated noun to be referentially definite and specific—just as the French pronominal affixes could be. In the example below, the morph corresponding to the meaning 'bed' is referred to by an anaphoric element which is part of the third word

in (32); as Baker (1996: 288) points out, the initial use of *-nakt-* is indefinite (and specific), but the incorporated pronoun *-ye-* is definite.

(32) Mohawk (Baker 1996: 288)

Thet´ʌre'	wa'-ke-nakt-a-hnínu-'.	Í-k-her-e'	Uwári
Yesterday	FACT-1sS-bed-0-buy-PUNC	0-1sS-think-IMPF	Mary

ʌ-ye-núhwe'-ne'
FUT-FsS-like-PUNC
'I bought a bed yesterday. I think Mary will like it (the bed)'

Baker (1996: 307) takes definite reference as criterial alongside the observation that incorporated nouns "can be modified by elements that appear outside the verb." For him, these phenomena argue against a lexicalist treatment of NI. However, we have seen that for Miller and Sag (1997), the ability to be independently referential is not criterial in establishing word-hood.

Baker (1996) states that NI is one of the central areas for investigating the lexicalist hypothesis, and he sets up a comparison of a version of a lexicalist theory against his own syntactic theory of the facts, which assumes Head Movement from the canonical argument position after the verb.[9] We can take the semantic argument first. I think that this is a non-argument: it is not at all clear that discourse reference is relevant to word-hood. There is a substantial literature which looks at how sub-events in word meanings can be modified (Levin 1993, Parsons 1990, Pustejovsky 1991). Gisborne (2010: 33) offers this example: *The submarine immediately sank for three hours.* In this case, the argument goes, there is no temporal conflict between *immediately* and *for three hours* so there must be two events in the meaning of *sank*, where each event is separately modified: one by *immediately* and one by *for three hours*. These events could be a 'going under' event and a 'staying under' event. An example of how a subpart of a word's meaning can be relevant to discourse reference would be *I reject the Thatcher$_i$ite view because she$_i$ didn't believe in society.*

[9] Ackema and Neeleman (2007) point out differences between Head Movement in NI, and more straightforward syntactic cases.

Discourse reference is necessarily conceptual. If a verb's meanings can interact with modifiers decompositionally, there is no reason to privilege discourse reference as a property of words. The research question is not, "How is it that some languages can have definite reference and discourse referents built into their verbs' meanings?" but "Why does English not include definite reference and discourse referents among the subparts of verbs' meanings that are independently accessible to the grammar?" This is a different reply to Baker from Anderson (2000, 2001), who suggests (following Rosen 1989, Di Sciullo and Williams 1987) that the verb with NI might well still be transitive, but that the argument is realized by a *pro* which is anaphoric to the incorporated nominal root. However it is entirely compatible with Miller and Sag's (1997) treatment of French pronominal affixes.

Anderson (2001) offers two independent arguments in favour of the lexicalist position. The first is that there is often non-systematic variation between the free form and the incorporated form (see also Osborne 1974). Quoting Mithun (1986: 876) he points out that in Mohawk "–*nahskw*- 'domestic animal' only appears incorporated while a semantically equivalent stem –*tshenv* appears only as an independent N[oun]".

The second is that the semantics of incorporated roots can be subject to lexical drift. This is particularly evident when the NI construction has a meaning which is different from its unincorporated analogue.

(33) Mohawk (Baker 1996; cited in Anderson 2000)

a. tu-s-a-yu-[a]t-háh-a-hkw-e'
dup-iter-fact-FsS-srfl-road-pick.up-punc
'She started her journey' (lit.: 'picked up the road')

b. #tu-s-a-yú-([a]te)-hkwe-e' ne oháha
Dup-iter-fact-fsS-srfl-pick.up-punc art road
'She picked up the road' (literal reading only)

The semantic arguments then are either irrelevant to an account of wordhood (as in the argument from discourse reference) or provide evidence in favour of the lexicalist approach. What about Baker's claim about apparently stranded modifiers? Anderson (2000) simply asserts that the material modifies the *pro* in the argument position. His difference from Baker (1996) is in his inventory of empty categories. Anderson (2000)

also points out, "In Mohawk, as in most languages, there is no overt correspondent of English *one*, so the object phrase in *I want a [new] one* consists of just the Adjective *new*.

However, although there is no inherent lexicalist reason to reject empty categories, the theories I mentioned at the beginning of this chapter all do. I reject empty categories because they are unlearnable, but no empty category is needed: the apparently stranded modifiers modify a subpart of the conceptual structure associated with the incorporated word, just as in my SINK example earlier. This is not very different from *he drinks his coffee black* where the adjunct of the Verb modifies an argument of the verb.

What then are the diagnostics which will decide whether NI is syntactic or morphological? The issue is important because a lot hangs on it. The debate goes back to Kroeber (1909) and Sapir (1911), with Kroeber arguing that there was no such thing as NI and Sapir providing a lexicalist analysis. More recently, Sadock (1980, 1985) has provided a syntactic account of incorporation phenomena in West Greenlandic and Baker (1988, 1996) has developed a syntactic approach which in turn provides a substantial part of the empirical and theoretical underpinnings of the theory of Head Movement.[10,11] On the other hand, Sapir (1911), Mithun (1986, 1988), Anderson (1992, 2000, 2001) Asudeh and Mikkelsen (2000) and others have argued that NI is built by morphological processes which are not part of syntax.

Baker (1996: 314-329) claims that there are three main arguments in favour of a syntactic solution to the problem of NI. First, he observes that the incorporated nominal element cannot be co-referential with the subject of the NI verb-form and argues that this is due to a Condition C binding effect—a referring expression cannot be locally bound—and that this follows from the incorporated nominal being extracted from the object position of the verb, and leaving behind a trace. However, Spencer (1995:00) extensively shows that this claim does not follow for other languages with NI, such as Chukchi which also show discourse reference effects with respect to the incorporated noun, and so are very similar to the Mohawk case. Moreover, Spencer shows that Baker's claim that adjuncts cannot incorporate does not apply to Chukchi, and effectively reduces Baker's theory to a set of parochial claims about Mohawk. It is also

[10] The phenomena Sadock (1980, 1985) discusses are arguably more like derivation than compounding (but see Spencer 1995) and some scholars such as Rosen (1989) have argued that they should not be included in discussions of incorporation. On the other hand, Sadock's approach brings the West Greenlandic phenomena in line with NI.

[11] Baker (2009) maintains that NI requires Head Movement despite other non-lexicalist theorists (Massam 2001, Van Geenhoven 2002) arguing that it is not necessary.

worth noting that the binding theory is itself subject to a number of theoretical challenges (Pollard and Sag 1992) and does not constitute a reliable diagnostic or analytic tool for other distributional facts.

Second, Baker shows that there are restrictions on questions: an interrogative form corresponding to WHO or WHAT cannot be extracted from object position when there is incorporation, which he accounts for by the argument that something else (the incorporated element) has already been extracted from object position. This argument does not wash in the terms of Principles and Parameters theory itself: Head Movement has a head position as its landing site; the landing site for WH operators in that theory is Spec,C. Moreover, we know that there are languages which permit the filling of both Spec,C and C: thus the literature on "Doubly-Filled Comps" in languages such as Belfast English (Henry 1995).

But Baker argues that his most important claim centres on the observation that there is no agreement with the incorporated (syntactic) argument in Noun Incorporation. As Anderson (2000) points out, this puts him in direct conflict with Postal (1979) who argues that there is agreement, but the absence of agreement is essential to Baker's position because he claims that it is not possible to have agreement with a trace, and his theory of NI requires there to be movement from the object position.

According to Baker (1996), the absence of agreement in NI structures is consistent with his claim that there is movement of a nominal element from the object position of the verb to the incorporated position. Baker (1996: 314-320) compares the structure in (34a) with that in (34b).

(34) a. The syntactic analysis of Baker (1988): $[N_i—V\ t_i]$
b. A lexicalist analysis with *pro*: [N—V pro]

The argument is that the syntactic account involves movement, leaving a trace behind, whereas the lexicalist account involves *pro*. Traces do not trigger agreement, so no agreement would be expected according to Baker's theory, whereas *pro* occupies the ordinary object slot, and so agreement would be expected according to the lexicalist account. Baker is not merely putting up a straw man: Rosen's lexicalist account calls for

pro-drop to account for the stranding facts (1989: 316), although the same is not true of Mithun (1984, 1986) and Mithun and Corbett (1999).

There is a complication. As Baker (1996: 315) points out, "Mohawk has no overt agreement morpheme for neuter objects". This is a particularly awkward wrinkle, because "[M]ost natural and spontaneous instances of NI n Mohawk generally involve neuter nouns [… which] has led to some confusion in the literature about whether I[ncorporated] N[ouns] are agreed with or not." Before I come to these points, I should make it clear that my own approach, with no empty categories, also predicts that there will be no agreement.

Once we reject empty categories, and if we at the same time treat Noun Incorporation as a straightforward case of compounding, there is nothing in the structure to trigger agreement. Therefore, if Baker is right, the absence of agreement is not evidence in favour of his movement-based approach. It is evidence against an approach which has the empty category *pro* as part of its analysis. As I said above, none of the lexicalist theories of syntax I mentioned in the introduction assumes empty categories, so the argument against a lexicalist theory that adopts *pro* is a strangely parochial argument.

Mohawk does occasionally incorporate animate nominal roots, and when it does they are also typically not accompanied by an agreement marker. I would have thought that the analysis was simple: it is not possible for a nominal root to be incorporated with its agreement markers in place, because we do not find inflection within words, and NI is a kind of compounding. But Baker (1996: 318) says that this argument will not work, because of data like those in (35), where there is an incorporated element and there is also doubling, in that it appears that the Object is simultaneously present with the incorporated nominal stem.

(35)	Ra-wir-a-núhwe'-s	thík^	(owirá'a)
	MsS-baby-0-like-HAB	that	baby
	'He likes that baby.'		

Baker (1996: 318) claims that this is the most important example in his book, because "it shows that verbs with incorporated nouns have a unique syntax […]" and "If the verb in (75) [=35, NG] were truly intransitive, then there would be no argument position to

licence the external NP in this sentence […] if the verb were a simple transitive, then it would have to agree with its object." The external NP in this case is the apparent object *thík^ (owirá'a)*, which Baker analyses as an adjunct. It is true that such data are difficult for the lexicalist analysis, because this means that the lexicalist account is obliged to have a theory of doubling which accounts for the absence of agreement in examples such as (35). One possibility of course is simply that (35) involves a structure such as *I like it, that baby.* Given the next item of evidence, it is certainly not enough to hang a whole theory off an example such as (35).

Baker (1996: 319) presents the examples in (36) as problematic for his theory. As he says, in these examples either an agreeing or a non-agreeing form is acceptable.

(36) a. Uwári ye(-ruwa)-kst^-hser-´^haw-e' ne rake-'níha
Mary FsS(/MsO)-old.person-NOM-carry-IMPF prt my-father
'Mary is holding my father'
b. Wa'-ke (-hi)-kst^-hser-áhset-e'
fact-1sS(/MsO)-old.person-NOM-hide-PUNC
'I hid the old person (the old man)'

He writes, "I would like to suggest that the material […] shows that *both* the lexical compounding analysis and the syntactic incorporation analysis are valid after all. The two structures often exist side by side in Mohawk". I am not sure how to address these facts. There is more relevant data however. Noun stems such as *-wir-* 'baby' can be incorporated, even though babies are not inanimate. In the example given, when *–wir-* is incorporated, there is no agreement. The examples are Baker's (1996: 316) examples (69) and (70).

(37) a. Shako-núhwe'-s (ne owirá'a)
MsS/3PO-likeHAB NE baby
'He likes them (babies).'
b. *Ra-núhwe'-s (ne owirá'a)
MsS-like-HAB NE baby
'He likes them (babies).'

(38) a. *?Shako-wir-a-núhwe'-s
MsS/3PO-baby-0-like-HAB
'He likes babies'
b. Ra-wir-a-núhwe'-s
MsS-baby-0-like-HAB
'He likes babies.'

Baker's argument hinges on examples like (37) and (38): the examples in (37) show that with the "ordinary transitive verb, feminine or masculine agreement is required when the direct object is understood to be a baby; null/neuter agreement is considered inappropriate". However the judgements are reversed in (38) when the noun is incorporated. This, Baker (1996: 316) claims is what is predicted by the theory in (34a), and it disconfirms the theory in (34b). But of course it is also entirely compatible with treating Mohawk incorporation as being a form of compounding with no empty category in the canonical object position.

4. The theoretical stance

In the three case studies I have discussed, I have argued for a lexicalist take on the relationship between the word and syntax, while pointing at a literature which often takes a very different position. In the case studies I have discussed, the theoretical proposal about the word and syntax has been discussed in the context of a larger theoretical package. We cannot just compare a generally lexicalist with a generally syntactic proposal: we need to discuss a specific syntactic proposal and compare it with an equally well worked out lexicalist proposal. The lexicalist position has turned out to be highly focused: for example, Miller and Sag's (1997) account of French pronominal affixes is concerned with the question of what the terminal node in a syntactic string is—and Miller and Sag argue that the terminal node is the word, not some subpart of a word. Therefore, they also take the view that subparts of words do not participate in syntax.

And this is the debate: are words, or sub-word elements (morphs, or morphemes), the primitives of syntax? My three case studies have been concerned with evaluating the evidence and exploring different proposals about this particular question: I haven't

been led to questions about the relationship between words and constructions, or words and their meanings. This last point has been particularly important in looking at pronominal affixes in §2 and the incorporation facts in §3. These assertions mean, of course, that I am not assuming that it is necessary (for the purposes of linguistic theory) to define "The Word" across a range of parameters, which is one of the objectives of Dixon and Aikhenvald (2002) and several of the papers therein.

This last point is important, because there are authors, such as Haspelmath (2011), who argue that the notion of the word should be dispensed with because it is too hard to define so that it is crosslinguistically valid and there is no simple convergent set of criteria for word-hood. In a sense there is also a general question about whether the word is the basic unit of analysis or whether the utterance is. Even this question has consequences for the theoretical status of the word. And so to the most important point: the word—the syntactic word in the case of this chapter—is a theoretical construct. Prototypical words might well exist in the wild, but when it comes to complex data involving boundaries between subparts of grammar, what a word is will depend on a host of intellectual choices made by the theorist. Hippisley, this volume, addresses similar issues.

There are three theoretical issues. What are the bounds on abstractness? What are the bounds on complexity? And where is the mess? I will take the last first. Mismatch is a kind of mess—arguably, most syntactic theorizing is about working out the limits on mismatch. A theory like Baker's (1996) reduces the mismatch, and therefore the mess, between syntax and semantics. His incorporated nominal stems can be in a one-to-one relationship with discourse referents. Baker puts the mess at the syntax-morphology interface: for Baker, what looks like a word is a syntagm. A theory like Miller and Sag's on the other hand locates the mess at the morphosyntax-semantics interface: an affixal part of a word can have an independent discourse referent, but their theory reduces the complexity at the morphology-syntax interface: the combinatorial units of syntax are words, and the problems of pronominal affixes being subject to affix ordering restrictions being able to undergo "clitic climbing" go away.

Abstractness is another issue. My account of the distribution of the passive participle in §2 relied on an abstract approach to morphology: I treated morphs as realizations of abstract feature clusters, so they do not have meaning in their own right, and I argued that this was the approach which best handled facts such as suppletion. A

less abstract theory of morphology would have morphemes—meaningful morphs which are not abstract realizations. But this would arguably add complexity, because it would mean that the lexicon would have to store a large number of homophonous morphemes. So, we might make a theory less abstract with a particular strategy, but that same strategy could make it more complex. Likewise, the decision about where the mess goes will have consequences for how abstract a theory is, and how complex. Baker's theory introduces the complexity of head movement. Miller and Sag's has the complexity of a mismatch between the word and discourse referents. Head movement gives us the abstractness of functional heads. Miller and Sag work with the abstractness of default inheritance and default overriding.

There are other dimensions of abstractness that I have not addressed. The version of a lexicalist analysis of NI that both Baker (1996) and Anderson (2000, 2001) present involves an unrealized pronoun, *pro*. Both Baker and Anderson assume that the debate about how to tackle NI reduces to a question of what the relevant empty category should be. While the debate can be framed in those terms, it is not obvious that it should be. After all, the theory proposed by Miller and Sag (1997) has no recourse to empty categories at all and it manages to account for apparent "clitic climbing" without clitics and without climbing.

References

Ackema, Peter and Ad Neeleman. 2002. Syntactic atomicity. *Journal of Comparative Germanic Syntax* 6, 93-128.

Ackema, Peter and Ad Neeleman. 2004. *Beyond Morphology: Interface Conditions on Word Formation*. Oxford: Oxford University Press.

Ackema, Peter and Ad Neeleman. 2010. The role of syntax and morphology in compounding. In Sergio Scalise and Irene Vogel, editors, *Cross Disciplinary Issues in Compounding*. Amsterdam: Benjamins, 21-36.

Alsina, Alex. 2009. The Prepositional Passive as Structure Sharing. In Miriam Butt and Tracy Holloway King, eds, *Proceedings of the LFG09 Conference*. CSLI Publications Stanford: CA. http://csli-publications.stanford.edu/

Anderson, Stephen R. (1982). Where's morphology? *Linguistic Inquiry* 13:571-612.

Anderson, Stephen R. 1992 *A-morphous Morphology*. Cambridge: Cambridge University Press.

Anderson, Stephen R. 2000. Some Lexicalist Remarks on Incorporation Phenomena. *Studia Grammatica* 45: 123-42.

Anderson, Stephen R. 2001. Lexicalism, Incorporated (or Incorporation, Lexicalized). *Proceedings of the 36th Annual Meeting of the Chicago Linguistics Society* pp. 13–34.

Anderson Stephen R. 2005 *Aspects of the Theory of Clitics*. Oxford: Oxford University Press.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge, MA: MIT Press.

Asudeh, Ash and Line Hove Mikkelsen. 2000. Incorporation in Danish: Implications for Interfaces. In Ronnie Cann, Claire Grover and Philip Miller, *Grammatical Interfaces in HPSG.* Stanford CA: CSLI Publications.

Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing.* Chicago: University of Chicago Press.

Baker, Mark C. 1996. *The Polysynthesis Parameter*. New York: Oxford University Press.

Baker, Mark C. 2009. Is Head Movement Still Needed for Noun Incorporation? The Case of Mapudungun. *Lingua* 119: 148-165.

Baltin, Mark and Paul Postal. 1996. More on Reanalysis Hypothesis. *Linguistic Inquiry* 27:127- 145.

Bonami, Olivier and Gilles Boyé. 2007. French Pronominal Clitics and the Design of Paradigm Function Morphology. In G. Booij, B. Fradin, A. Ralli and S. Scalise, *Online Proceedings of the Fifth Mediterranean Morphology Meeting* 291-322.

Bresnan, Joan 1978. A realistic transformational grammar. In Morris Halle, Joan Bresnan, and George A. Miller, eds., *Linguistic theory and psychological reality*. Cambridge, MA: MIT Press. 1-59.

Bresnan, Joan. 1982. The passive in lexical theory. In Joan Bresnan (editor), *The Mental Representation of Grammatical Relations*, Cambridge, MA: MIT Press. 3-86

Bresnan, Joan. 2001. *Lexical-Functional Syntax.* Oxford: Blackwell Publishers.

Brown, Dunstan and Andrew Hippisley. 2012. *Network Morphology.* Cambridge: Cambridge University Press.

Bruening, Benjamin. 2012. Word Formation is Syntactic: Adjectival Passives in English. To appear in *Natural Language and Linguistic Theory.* Downloaded from http://udel.edu/~bruening/Downloads/AdjectivalPass3.pdf 15 February 2013.

Chomsky, Noam. 1957. *Syntactic Structures.* The Hague: Mouton.

Chomsky, Noam. 1970. Remarks on nominalization. In R. Jacobs and P. Rosenbaum eds., *Readings in English Transformational Grammar.* Waltham, MA: Ginn. 184-221.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Collins, Christopher 2005. A Smuggling Approach to the Passive in English. *Syntax* 8: 81-120.

DiSciullo, Ann Maria and Edwin Williams. 1987. *On the Definition of Word*. Cambridge, MA: MIT Press.

Dixon, R.M.W. and Alexandra Aikhenvald. 2002. *Word: a cross-linguistic typology.* Cambridge: Cambridge University Press.

Embick, David and Rolf Noyer. 2007. Distributed Morphology and the Syntax/Morphology Interface. In Gillian Ramchand and Charles Reiss, eds., *The Oxford Handbook of Linguistic Interfaces* Oxford: Oxford University Press, 289–324.

Embick, David and Alec Marantz. 2008. Architecture and blocking. *Linguistic Inquiry* 39, 1-53.

Friedin, Robert. 1975. The analysis of passives. *Language* 51. 385-405.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum, and Ivan Sag. 1985. *Generalized Phrase Structure Grammar.* Oxford: Basil Blackwell.

Giegerich, Heinz J. 2009. Compounding and Lexicalism. In Lieber, R., and Stekauer, P. eds., *The Oxford Handbook of Compounding.* Oxford: Oxford University Press . 178-200.

Gisborne, Nikolas. 2008. Dependencies are Constructions. In Graeme Trousdale and Nikolas Gisborne, eds., *Constructional Approaches to English Grammar*. Berlin: Mouton de Gruyter. 219-255.

Gisborne, Nikolas. 2010. *The Event Structure of Perception Verbs.* Oxford: Oxford University Press.

Halle Morris. 1973. Prolegomena to a Theory of Word Formation. *Linguistic Inquiry* 4, 3-16.

Halle, Morris and Alec Marantz. 1993. Distributed Morphology and the Pieces of Inflection. In Kenneth Hale and S. Jay Keyser, eds., *The View From Building 20*. Cambridge, MA: MIT Press, 111-176.

Haspelmath, Martin. 2011. The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia Linguistica* 45, 31-80.

Henry, Alison. 1995. *Belfast English and Standard English.* New York: Oxford University Press.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge: Cambridge University Press.

Hudson, Richard. 1990. *English Word Grammar.* Oxford: Basil Blackwell.

Hudson, Richard. 2003. Gerunds Without Phrase Structure. *Natural Language and Linguistic Theory* 21, 579-615.

Hudson, Richard. 2007. *Language Networks: The New Word Grammar*. Oxford: Oxford University Press.

Jaeggli, Osvaldo. 1982. *Topics in Romance Syntax.* Dordrecht: Foris.

Kayne, Richard S. 1969. *The Transformational Cycle in French Syntax.* Doctoral Dissertation, MIT.

Kayne, Richard S. 1975. *French Syntax: The Transformational Cycle*. Cambridge, MA: MIT Press.

Kayne, Richard S. 1991. Romance Clitics, Verb Movement and PRO. *Linguistic Inquiry*, 22, 647-686.

Kroeber, A. L. 1909. Noun Incorporation in American Languages. XVI Internationaler Amerikanisten-Kongress, 569-576.

Lapointe, Steven G. 1980. *A Theory of Grammatical Agreement.* Doctoral Dissertation, University of Massachusetts at Amherst.

Levin, Beth. 1993. *Verb Classes and Alternations.* Chicago: University of Chicago Press.

Levin, Beth and Malka Rappaport. 1986. The Formation of Adjectival Passives. *Linguistics Inquiry* 17, 623-661.

Lieber, Rochelle and Sergio Scalise. 2007. The Lexical Integrity Hypothesis in a New Theoretical Universe. In G. Booij, et al, eds., *On-line proceedings of the Fifth Mediterranean Morphology Meeting* Fréjus 15-18 September 2005, University of bologna, 2007. http://mmm.lingue.unibo.it/

Lødrup, Helge (1991) The Norwegian Pseudo Passive as Structure Sharing. *Working Papers in Scandinavian Syntax* 47, 118-29.

Malouf , Robert 1998. *Mixed Categories in The Hierarchical Lexicon.* Doctoral Dissertation, Stanford University.

Matthews, P.H. 1991. *Morphology*. Cambridge: Cambridge University Press.

McCawley, James D. 1998. *The Syntactic Phenomena of English*. Chicago: Chicago University Press.

Miller, Philip. 1992. *Clitics and Constituents in Phrase Structure Grammar*. Doctoral Dissertation, Utrecht University.

Miller, Philip and Ivan Sag. 1997. French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15, 573-639.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60, 847-94.

Mithun, Marianne. 1986. On the nature of noun incorporation. *Language* 62, 32-37.

Mithun, Marianne. 2000. Incorporation. *Morphologie/Morphology: A Handbook on Inflection and Word Formation: Ein internationales Handbuch zur Flexion und Wortbildung.* Geert Booij, Christian Lehmann, Joachim Mudgan, eds., Berlin: Walter de Gruyter, 916-28.

Mithun, Marianne and Greville G. Corbett. 1999. The effect of noun incorporation on argument structure. In Lunella Mereu ed., *The Boundaries of Morphology and Syntax.* Amsterdam: John Benjamins, 49-71.

Osborne, C. R. 1974. *The Tiwi Language*. Australian Aboriginal Studies, 55. Canberra: Australian Institute of Aboriginal Studies.

Parsons, Terence. 1990. *Events in the Semantics of English.* Cambridge, MA: MIT Press.

Perlmutter, David M. 1970. Surface Structure Constraints in Syntax. *Linguistic Inquiry* 1, 187–255.

Pollard, Carl, and Ivan A. Sag. 1992. Anaphors in English and the scope of binding theory. *Linguistic Inquiry* 23, 261-303.

Pollard, Carl and Ivan Sag. 1994. *Head-driven Phrase Structure Grammar.* Chicago: University of Chicago Press.

Postal Paul M. 1979 *Some Syntactic Rules in Mohawk*. New York, Garland.

Postal, Paul M. 1985. *Studies of Passive Clauses.* Stony Brook: SUNY Press.

Postal, Paul M. 1996. A Glance at French Pseudopassives. In C. Burgess, K. Dziwirek and D. B. Gerdts eds., *Grammatical Relations, Theoretical Approaches to Empirical Questions*. Standford, CA: CSLI Publications.

Pullum, Geoffrey K. 2011. The passive in English. Language Log post of 24 January 2011. Available at http://languagelog.ldc.upenn.edu/nll/?p=2922 Downloaded on 12 March 2013.

Pustejovsky, James. 1990. The syntax of event structure. *Cognition* 41, 47-81.

Rivas, Alberto M. 1977. A theory of clitics. Doctoral dissertation, MIT.

Rosen, Sara Thomas. 1989. Two Types of Noun Incorporation: A Lexical Analysis. *Language* 65, 294-317.

Sadock, Jerrold M. 1980. Noun Incorporation in Greenlandic: A Case of Syntactic Word-Formation. *Language* 57, 300-19.

Sadock, Jerrold M. 1985. Autolexical Syntax: A Theory of Noun Incorporation and Similar Phenomena. *Natural Language and Linguistic Theory* 3, 379-441.

Sadock, Jerrold M. 1991. *Autolexical Syntax: A Theory of Parallel Grammatical Representations.* Chicago: University of Chicago Press.

Sapir, Edward. 1911. The Problem of Noun Incorporation in American Languages. *American Anthropologist* 13, 250-282.

Seiter, William J. 1980. *Studies in Niuean Syntax.* New York: Garland Publishing Company.

Spencer, Andrew. 1995. Incorporation in Chukchi. *Language* 71, 439-89.

Spencer, Andrew, and Ana Luis. 2012. *Clitics: An Introduction*. Cambridge: Cambridge University Press.

Sportiche, Dominque. 1996. Clitic Constructions. in J. Rooryck and L. Zaring eds., *Phrase Structure and The Lexicon*. Bloomington, Indiana: IULC Press, 213–276.

Stump, Gregory N. 2001. *Inflectional Morphology: A Theory of Paradigm Structure.* Cambridge: Cambridge University Press.

Wasow, Thomas. 1977. Transformations and the lexicon. In Peter Culicover, Thomas Wasow, and Joan Bresnan, eds., *Formal Syntax*. New York: Academic Press. 327-360.

Woodbury, Hanni. 1975. Onondaga Noun Incorporation: Some Notes on the Interdependence of Syntax and Semantics. *International Journal of American Linguistics* 41, 10-20.

Zwicky, Arnold M. 1977. On clitics. Bloomington, IN: Indiana University Linguistics Club.

Zwicky, Arnold M. 1995. What is a clitic? In Joel Nevis, Brian Joseph, Dieter Wanner, and Arnold M. Zwicky, eds., *Clitics: A Comprehensive Bibliography 1892–1991*. XII-XX.

Zwicky, Arnold M. and Geoffrey K. Pullum. 1983. Cliticization vs. Inflection: English N'T. *Language* 59, 502-13.